

CLAIMS

1. An arrangement allowing multi-modal access of content over a
5 global data communications network, e.g. Internet, comprising a
mobile station (MS), with a user agent, a proxy server, and a
telephony platform,
c h a r a c t e r i z e d i n
that the mobile station is a dual mode station supporting
10 concurrent voice and data sessions,
that the proxy server comprises an enhanced functionality for
supporting voice browsing,
that the telephony platform comprises an Automatic Speech
Recognizer (ASR) and a block for converting text messages to
15 speech,
that said enhanced proxy interfaces the Automatic Speech
Recognizer of the Telephony Platform, that key elements (e.g.
text, words phrases) are predefined and indicated in the
(original) web content,
20 and in that when the enhanced proxy server recognizes/extracts
said key elements (using predefined rules) it triggers voice
browsing, such that an arbitrary web content (page) can be
accessed by voice commands without requiring conversion of the
web content.
- 25
2. An arrangement according to claim 1,
c h a r a c t e r i z e d i n
that multi-modal browsing is implemented.
- 30
3. An arrangement according to claim 1 or 2,
c h a r a c t e r i z e d i n
that the enhanced proxy server parses an accessed web content
with regard to said key elements.

4. An arrangement according to any one of claims 1-3,
c h a r a c t e r i z e d i n
that the accessed web content is browsed by means of key strokes,
5 mouse clicks or similar.
5. An arrangement according to any one of the preceding claims,
c h a r a c t e r i z e d i n
that it allows for voice-based access of any tag based content,
10 e.g. HTML/XHTML web content.
6. An arrangement according to any one of the preceding claims,
c h a r a c t e r i z e d i n
that the user of the mobile station uses a key element indicated
15 in the web content to select a specific hyperlink.
7. An arrangement according to any one of the preceding claims,
c h a r a c t e r i z e d i n
that the voice browsing functionality of the enhanced proxy
20 server implements keyword spotting.
8. An arrangement according to claim 8,
c h a r a c t e r i z e d i n
that the enhanced proxy server interfaces with the Automatic
25 Speech Recognizer which comprises a medium size vocabulary
speech recognizer.
9. An arrangement according to any one of the preceding claims,
c h a r a c t e r i z e d i n
30 that the predefined rules for voice key element extraction are
syntactic rules.
10. An arrangement according to any one of claims 1-8,

c h a r a c t e r i z e d i n
that the predefined rules for voice key element extractions are
simple rules, e.g. relating to selection of a unique keyword in
the name of a hyperlink.

5

11. An arrangement according to any one of claims 1-8,
c h a r a c t e r i z e d i n
that the predefined rules for voice key element extraction are
numeric rules numbering hyperlinks in a content or similar.

10

12. An arrangement according to any one of the preceding claims,
c h a r a c t e r i z e d i n
that the enhanced proxy server forwards text prompts, to the Text
to Speech block in the telephony platform, wherein the text
15 messages are converted to speech and forwarded to the user over
the voice channel set up by the enhanced proxy server.

13. An arrangement according to any one of the preceding claims,
c h a r a c t e r i z e d i n

20 that between the conventional browser in the user agent and the
speech browser in the enhanced proxy server a synchronization
engine is provided.

14. An arrangement according to claim 13,

25 c h a r a c t e r i z e d i n

that the enhanced proxy server comprises a pushing mechanism for
making the MS user agent refresh indicated, fetched content.

15. An arrangement according to claim 14,

30 c h a r a c t e r i z e d i n

that a semaphore object is introduced into the content returned
to the enhanced server for indicating activation or not of
content refresh.

16. An arrangement according to any one of the preceding claims,
c h a r a c t e r i z e d i n
that a connection is established between the enhanced proxy
5 server and the Automatic Speech Recognizer of the telephony
platform for specifying and identifying a called application to
be accessed.

17. An arrangement according to claim 16,
10 c h a r a c t e r i z e d i n
that the enhanced proxy server comprises a number of subscriber
(end user) records, and in that for each subscriber for which
voice browsing should be supported, means for indication of
voice browsing activation, optional key element (word) for
15 triggering voice browsing or optional hyperlink name, for
insertion in accessed web page/content, and which, when
selected, provides for establishment of a voice channel between
the Automatic Speech Register and the mobile station.

20 18. An arrangement according to claim 16 or 17,
c h a r a c t e r i z e d i n
that if voice browsing is activated, the access request is
forwarded from the enhanced proxy server to the relevant
Application Service Provider, which returns the requested
25 page/content to the enhanced proxy server, and in that said
enhanced proxy server comprises parsing and analyzing means for
finding and indicating key elements (words), before forwarding
the content/page as modified to the mobile station.

30 19. An arrangement according to any one of the preceding claims,
c h a r a c t e r i z e d i n
that a request for voice browsing has to include at least a voice
browsing session ID and MSISDN of the user station.

20. An arrangement according to claim 19,
c h a r a c t e r i z e d i n
that for a user authenticated by the enhanced proxy server, a
5 voice channel is established, concurrent with a data session
channel, between the Automatic Speech Register and the mobile
station.

21. An arrangement according to any one of claims 18-20,
10 c h a r a c t e r i z e d i n
that keywords as recognized in voice commands from the end user
are provided to the enhanced proxy server, and in that the
enhanced proxy server comprises matching means for matching
recognized voice commands with stored key elements/words, for
15 finding the relevant link on which to send a request to the
Application Service Provider, and in that the requested content,
upon reception in the enhanced proxy server, is parsed, analyzed
and pushed to the user agent.

22. An arrangement according to claim 12,
c h a r a c t e r i z e d i n
that for synchronization between the user agent of the mobile
station and the enhanced proxy server, a client semaphore object
is introduced, (by the enhanced proxy server) into the original
25 content ((X)HTML) of which the original copy is stored in said
server, and activated (ON) when voice browsed content is to be
pushed to be mobile station.

23. An arrangement according to claim 22,
30 c h a r a c t e r i z e d i n
that the client semaphore object is periodically updated with the
value of the semaphore object in the enhanced proxy server.

24. An arrangement according to claim 23,
c h a r a c t e r i z e d i n
that in the user agent (client) a script downloaded with original
content continuously checks the client semaphore object to
5 establish if a content refresh is required, and in that, in the
enhanced proxy server, a script is used to activate the proxy
semaphore object (ON).

25. An arrangement according to claim 23 or 24,
10 c h a r a c t e r i z e d i n
that the client semaphore object is created using a WML script
variable, fetched from the enhanced proxy server, and in that in
the enhanced proxy server a first and a second version of said
script is stored, the first version comprising a script for
15 semaphore activation (ON), the second version comprising a
script indications semaphore inactive.

26. A method for enabling/providing concurrent multi-modal
access of global datacommunication networks, e.g. Internet
20 content (a page etc.) from a dual mode mobile station,
c h a r a c t e r i z e d i n
that it comprises the steps of:
- providing a proxy server with an enhanced functionality for
voice browsing,
25 - defining rules for keyword extraction from a browsed content
and keywords/key elements,
- indicating the keywords in the original content,
- based on indication of keywords, end user selecting a keyword
to select a specific link/hyperlink such that an arbitrary web
30 content/page can be accessed by voice without requiring
conversion of the original content.

27. A method for providing concurrent multi-modal access of an Internet content (e.g. a page etc.) from a dual mode mobile station,

c h a r a c t e r i z e d i n

5 that it comprises the steps of:

- providing an enhanced functionality proxy server supporting voice browsing,
- establishing a connection between the enhanced proxy server and a telephony platform with an Automatic Speech Register (ASR),
- 10 - establishing/defining key elements (words) to use at voice browsing,
- establishing if voice browsing is to be active and supported, if yes,
- 15 - setting up a voice channel between the mobile station and the Automatic Speech Register (based on user profile),
- forwarding a request to the concerned application service provider,
- parsing content and analyzing paragraphs in the content/web
- 20 page to find key elements,
- modifying, in the enhanced proxy, the content by changing tag attributes to make key elements identifiable to the user,
- sending the content modified as in the preceding step to the mobile station,
- 25 - opening a voice browsing session,
- opening a voice channel/concurrent with data session channel,
- matching, in the enhanced proxy server, keywords recognized in user's voice command with predefined and selected keywords to establish which link to use for sending a get request to the
- 30 relevant application service provider,
- processing and pushing the content received from the application service provider to the user agent.